SOLUTIONS.

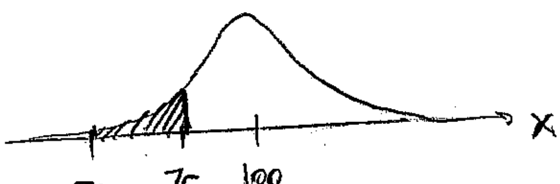
Alex Barnett
2/21/06.

Math 50: Midterm 2

65 minutes, 70 points. No algebra-capable calculators. Try to use your calculator minimally—you barely need it. Show working/reaonsing, since only that way could you get partial credit.

1. [10 points] IQ (the supposed 'intelligence quotient') is an integer scale designed to be normally-distributed in the population, with $\mu = 100$ and $\sigma = 15$.

   (a) What fraction of the population is then required to be a 'moron' (a technical term, defined by $50 < \text{IQ} < 75$) ?



Use $Z = \frac{X-\mu}{\sigma}$ where $X = $ IQ.

continuity correction makes v. little difference here since $\sigma = 15 \gg 1$.

$z = \frac{75-100}{15} = -1.67$

$z = \frac{50-100}{15} = -3.33$

fraction $= F_Z(-1.67) - F_Z(-3.33)$ so can use table

$= F_Z(+3.33) - F_Z(+1.67)$

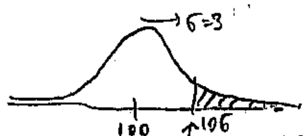
looking up 'far right tail'

$= (1 - 0.00048) - 0.9525 = 0.0470 = 4.7\%$

   (b) What is the chances that the average IQ of a random sample of size 25 of the population has an IQ of at least 106? [Hint: IQ is an integer quantity; but you will not lose much for ignoring this]

$\text{Var}(\bar{X}) = \frac{1}{n} \text{Var}(X)$ so $\sigma(\bar{X}) = \frac{1}{\sqrt{n}} \sigma = \frac{15}{\sqrt{25}} = 3$.

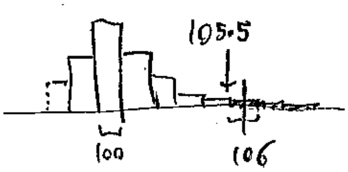
If IQ were continuous quantity,



$z = \frac{106-100}{3} = 2$

$\text{prob}(\bar{X} \geq 106) = 1 - F_Z(2) = 0.0228$

But since IQ is discrete (integer) and σ is small, use continuity correction:

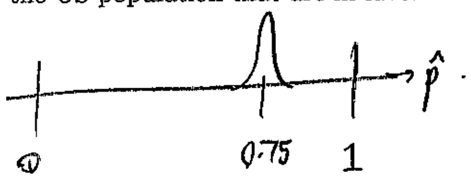


$Z = \frac{105.5-100}{3} = \frac{11}{6}$ so

$\text{prob}(\bar{X} \geq 106) = 1 - F_Z(\frac{11}{6}) = \underline{0.0336}$.

2. [12 points] 1000 people randomly sampled from the US population are given a survey asking if they are in favor of gay marriage.

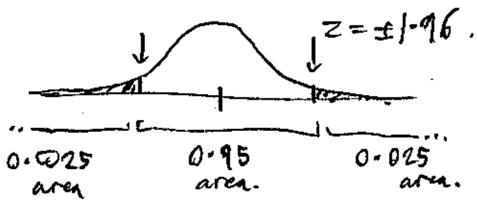
(a) Suppose 750 of the 1000 are in favor. Construct a 95% confidence interval on p, the fraction of the US population that are in favor.



$\hat{p} = \frac{X}{n}$ $\qquad n = 10^3$

where $p = \frac{k}{n} = \frac{3}{4}$ has been substituted (a hack).

$$\sigma = \sqrt{Var(\hat{p})} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{3/16}{1000}} = 0.01377$$



Interval is $\left[\frac{3}{4} - 1.96\sigma, \frac{3}{4} + 1.96\sigma\right]$

$= [0.723, 0.777]$

(b) Suppose p is unknown and you want to design a survey to estimate p with a *margin of error* of 3%. What is the minimum number of people you need to survey?

Worst-case $p = 1/2$ so $\sigma = \sqrt{\frac{1}{4n}} = 0.03$

meaning $\pm 3\%$, ie $\sigma = 0.03$ the std. dev. of estimator $\hat{p} = \frac{X}{n}$.

$\Rightarrow \quad n = \frac{1}{4(0.03)^2}$

$= 277.8$

So n must be at least $\underline{278}$.

3. [23 points] Data are drawn from the model pdf

$$f_Y(y;\theta) = 2y/\theta^2 \quad \text{for} \quad 0 < y < \theta, \quad \text{zero otherwise.}$$

Given samples $\{y_1, \ldots, y_n\}$, we wish to estimate the parameter θ.

4 (a) Find the Method of Moments estimator $\hat{\theta}$.

1 unknown param ⇒ use only 1 moment

sample 1st moment

$$\text{Equate} \quad E(Y) = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$\int_0^\theta f_Y(y;\theta)\, y\, dy = \frac{2}{\theta^2}\int_0^\theta y^2\, dy = \frac{2}{\theta^2}\cdot\frac{\theta^3}{3} = \frac{2}{3}\theta.$$

$$\text{so} \quad \hat{\theta} = \frac{3}{2}\cdot\frac{1}{n}\sum_{i=1}^{n} y_i = \frac{3}{2}\overline{Y}$$

3 (b) Is this estimator unbiased? (Prove your answer)

Does $E(\hat{\theta}) = \theta$ the "true" underlying param. value?

$$E(\hat{\theta}) = \frac{3}{2}E(\overline{Y}) = \frac{3}{2}E(Y) = \frac{3}{2}\cdot\frac{2}{3}\theta = \theta \quad \checkmark$$

since averaging doesn't change $E(\cdot)$

worked out above

unbiased.

5 (c) What is the efficiency of this estimator, Var$(\hat{\theta})$?

$$\text{Var}(\hat{\theta}) = \left(\frac{3}{2}\right)^2 \text{Var}(\overline{Y}) = \left(\frac{3}{2}\right)^2 \frac{1}{n}\text{Var}(Y)$$

with
$$\text{Var}(Y) = E(Y^2) - E(Y)^2$$
$$= \frac{2}{\theta^2}\int_0^\theta y^2 \cdot y\, dy - \left(\frac{2}{3}\theta\right)^2 = \frac{2}{\theta^2}\cdot\frac{\theta^4}{4} - \frac{4}{9}\theta^2$$
$$= \frac{\theta^2}{18}.$$

$$\text{so} \quad \text{Var}(\hat{\theta}) = \frac{9}{4}\cdot\frac{1}{n}\cdot\frac{\theta^2}{18} = \boxed{\frac{\theta^2}{8n}}$$

4. [11 points] A coin of unknown bias $0 \le p \le 1$ is flipped 3 times and gives the data: heads, tails, heads.

(a) Assuming an uninformative prior, compute the (correctly-normalized) posterior pdf on p given this data.

$$\text{Likelihood} \quad L(p) = c\, \underbrace{p^2}_{\text{2 heads}} \underbrace{(1-p)^1}_{\text{1 tail.}}$$

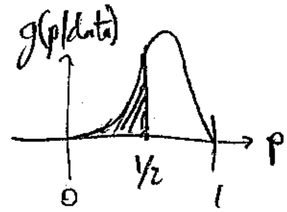
prior is constant

$$\Rightarrow \text{posterior} \quad g(p \mid \text{data}) = c \cdot p^2 (1-p)$$

What is c? It's beta pdf with $r = 3$, $s = 2$

$$\Rightarrow c = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)} = \frac{\Gamma(5)}{\Gamma(3)\Gamma(2)} = \frac{4!}{2!\,1!} = 12$$

(b) Given this posterior, compute $P(p \le 1/2)$, that is, the Bayesian answer to the question, "what is the chance that the coin is biased in the tails direction?"



$$P(p \le 1/2 \mid \text{data}) = \int_0^{1/2} g(p \mid \text{data})\, dp$$

$$= \int_0^{1/2} 12\, p^2 (1-p)\, dp$$

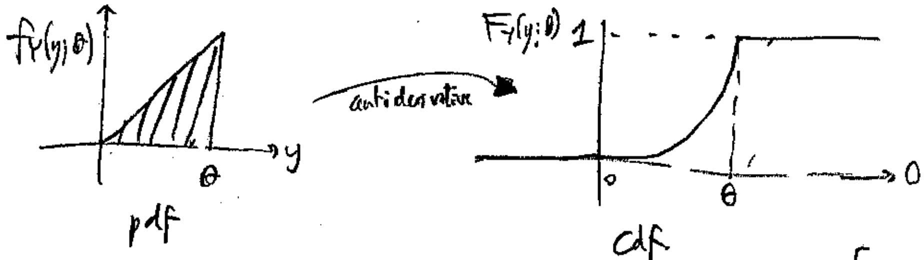
$$= 12 \left[\frac{p^3}{3} - \frac{p^4}{4} \right]_0^{1/2}$$

$$= 12 \cdot \frac{1}{16} \underbrace{\left(\frac{2}{3} - \frac{1}{4} \right)}_{\frac{8}{12} - \frac{3}{12} = \frac{5}{12}}$$

$$= \underline{\frac{5}{16}}.$$

interesting that not $\frac{1}{3}$ as might guess from data.

5 (d) As with the uniform pdf, the Maximum Likelihood estimator is $\hat{\theta}_{ML} = Y_{max}$. What is the bias of this estimator? If needed, suggest a fix which makes it unbiased.



$$F_Y(y;\theta) = \frac{2}{\theta^2}\int_0^y y'\,dy' = \frac{y^2}{\theta^2}$$

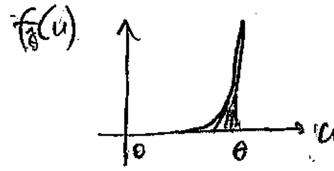
$$f_{Y_{max}}(y) = n\left(\frac{y^2}{\theta^2}\right)^{n-1}\frac{2y}{\theta^2}$$

$$E(\hat{\theta}) = E(Y_{max}) = \int_0^\theta y\, f_{Y_{max}}(y)\, dy$$

$$= \frac{2n}{\theta^{2n}} \cdot \underbrace{\int_0^\theta y\, y^{2n-2}\, y\, dy}_{\text{gives } \frac{\theta^{2n+1}}{2n+1}} = \frac{2n}{2n+1}\theta \quad \text{biased.}$$

→ Unbiased version $\hat{\theta} = \frac{2n+1}{2n} Y_{max}$

4 (e) Prove whether the estimator $\hat{\theta}_{ML}$ is consistent or not. → (the unfixed ML one, although both are consistent).



← pdf of estimator is $f_{Y_{max}}(u) = n\left(\frac{u^2}{\theta^2}\right)^{n-1}\frac{2u}{\theta^2} = \frac{2n}{\theta^{2n}} \cdot u^{2n-1}$ (or you may use y as variable)

Consistent if $\underbrace{p(\theta - \varepsilon \le \hat{\theta} \le \theta)} \to 1$, as $n \to \infty$.

$$\int_{\theta-\varepsilon}^{\theta} f_{Y_{max}}(u)\, du = \frac{2n}{\theta^{2n}}\left[\frac{u^{2n}}{2n}\right]_{\theta-\varepsilon}^{\theta} = \frac{1}{\theta^{2n}}\left[\theta^{2n} - (\theta-\varepsilon)^{2n}\right] = 1 - \underbrace{\left(1-\frac{\varepsilon}{\theta}\right)^{2n}}_{\text{vanishes as } n\to\infty}$$

Therefore our probability → 1 as $n \to \infty$.

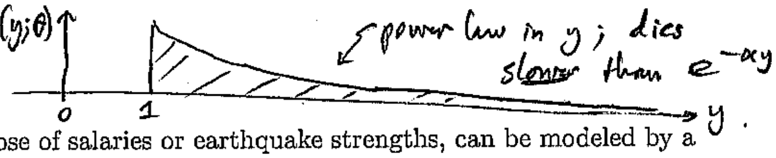
NB: You instead could have shown (for unbiased version) variance $= \frac{\text{const}}{n^2} \to 0$. (long calculation).

2 (f) Give an example of an estimator which is *not* consistent (either for the above pdf, or any pdf of your choosing).

eg. $\hat{\theta} = \frac{2}{3}y_1$, the first sample (ignoring all the n-1 rest samples).

or $\hat{\theta} = \bar{Y}$ from Cauchy pdf (HW6).

Note: an <u>estimator</u> is a <u>function of the data only</u>, so you can't use θ in your example estimator!



5. [14 points] Some distributions, such as those of salaries or earthquake strengths, can be modeled by a power-law pdf with parameter $\theta > 0$, thus

$$f_Y(y;\theta) = \theta y^{-1-\theta}, \quad y \geq 1, \quad \text{zero otherwise.}$$

5 (a) Given n samples $\{y_i\}$, find the ML estimator. [Hint: $y^{-\theta} = e^{-\theta \ln y}$]

$$L(\theta) = \prod_{i=1}^{n} f_Y(y_i;\theta) = \theta^n \prod_i y_i^{-(1+\theta)} \rightarrow e^{-(1+\theta)\ln y_i}$$

take log:

$$\ln L = n \ln\theta + -(1+\theta)\sum_{i=1}^{n} \ln y_i$$

$$\frac{\partial}{\partial\theta} \ln L = \frac{n}{\theta} - \sum_i \ln y_i \longrightarrow \text{set to zero, to maximize } L.$$

$$\Rightarrow \hat{\theta} = \frac{n}{\sum_i \ln y_i} \qquad \text{ML estimate.}$$

5 (b) Find the Cramér-Rao bound on the variance of any estimator for θ. Be sure to state whether it's a lower or upper bound.

$$\ln f_Y(Y;\theta) = \ln\left(\theta Y^{-1-\theta}\right) = \ln\theta - (1+\theta)\ln Y$$

derivs.

$$\frac{\partial}{\partial\theta} \ln f_Y = \frac{1}{\theta} - \ln Y$$

$$\frac{\partial^2}{\partial\theta^2} \ln f_Y = \frac{-1}{\theta^2}, \text{ is } \underline{\text{indep. of } Y}, \text{ so } E\left[\frac{\partial^2}{\partial\theta^2} \ln f_Y\right] = \frac{-1}{\theta^2}, \text{ too!}$$

$$\Rightarrow CR: \quad \mathrm{Var}(\hat{\theta}) \geq \frac{-1}{nE\left[\frac{\partial^2}{\partial\theta^2}\ln f_Y\right]} = \frac{-1}{n(-1/\theta^2)} = \boxed{\frac{\theta^2}{n}} \quad \text{lower bound.}$$

4 (c) What pdf is the *conjugate prior* for this power-law pdf? (you must show *why*)

Likelihood has form, ~~as~~ as func of θ, $L(\theta) = \underbrace{\left(\prod_i y_i\right)}_{\theta\text{-indep. constant.}} \cdot \underbrace{\theta^n e^{-\theta \sum_i \ln y_i}}_{\theta^{r-1} e^{-\lambda\theta}}$

this is hard to make mental switch (y = const now!)

this is $\underline{\text{gamma pdf}}$ with params $\begin{cases} r = n+1 \\ \lambda = \sum_i \ln y_i \end{cases}$

Therefore if prior is also gamma, posterior is again gamma (same family), $\boxed{\text{Gamma pdf is conjugate prior.}}$